Hon. Robert J. Bryan

UNITED STATES DISTRICT COURT
WESTERN DISTRICT OF WASHINGTON

UNITED STATES OF AMERICA, *Plaintiff,* v. KING COUNTY, WASHINGTON, et al., *Defendants.*	CASE NO. 2:20-cv-203 **[DRAFT] AGREEMENT REGARDING DISCOVERY OF ELECTRONICALLY STORED INFORMATION AND [PROPOSED] ORDER**

The parties hereby stipulate to the following provisions regarding the discovery of electronically stored information ("ESI") in this matter:

A.   General Principles

1.   An attorney's zealous representation of a client is not compromised by conducting discovery in a cooperative manner. The failure of counsel or the parties to litigation to cooperate in facilitating and reasonably limiting discovery requests and responses raises litigation costs and contributes to the risk of sanctions.

2.   As provided in LCR 26(f), the proportionality standard set forth in Fed. R. Civ. P. 26(b)(1) must be applied in each case when formulating a discovery plan. To further the application of the proportionality standard in discovery, requests for production of ESI and related responses should be reasonably targeted, clear, and as specific as possible.

1   **B.   ESI Disclosures**

2   Within 30 days of entry of this Order, or at a later time if agreed to by the parties in writing,

3   each party shall disclose:

4   1.   Custodians. The five custodians most likely to have discoverable ESI in their

5   possession, custody, or control. The custodians shall be identified by name, title, connection to

6   the instant litigation, and the type of the information under the custodian's control. The parties

7   will confer in good faith if, after review of documents from these initial custodians, one or both

8   parties believe additional custodial collections are necessary.

9   2.   Non-Custodial Data Sources. A list of non-custodial data sources (*e.g.*, shared

10   drives, servers), if any, likely to contain discoverable ESI.

11   3.   Third-Party Data Sources. A list of third-party data sources, if any, likely to

12   contain discoverable ESI (*e.g.*, third-party email providers, mobile device providers, cloud

13   storage) and, for each such source, the extent to which a party is (or is not) able to preserve

14   information stored in the third-party data source.

15   4.   Inaccessible Data. A list of data sources, if any, likely to contain discoverable ESI

16   (by type, date, custodian, electronic system or other criteria sufficient to specifically identify the

17   data source) that a party asserts is not reasonably accessible under Fed. R. Civ. P. 26(b)(2)(B).

18   Data sources set forth in Section (D)(3) of this Agreement are exempt from this requirement.

19   **C.   ESI Discovery Procedures**

20   1.   On-site inspection of electronic media. Such an inspection shall not be required

21   absent a demonstration by the requesting party of specific need and good cause or by written

22   agreement of the parties.

23

24

25   [DRAFT] AGREEMENT REGARDING DISCOVERY OF ELECTRONICALLY
     STORED INFORMATION AND [PROPOSED] ORDER   PAGE - 2
26   (Case No. 2:20-CV-203)

1   2.   <u>Search methodology.</u> The parties shall timely confer to attempt to reach agreement

2 on appropriate search terms and queries, file type and date restrictions, data sources (including

3 custodians), and other appropriate computer- or technology-aided methodologies, before any such

4 effort is undertaken. The parties shall continue to cooperate in revising the appropriateness of the

5 search methodology.

6   a.   Prior to running searches:

7   i.   The producing party shall disclose the data sources (including

8 custodians), search terms and queries, any file type and date restrictions, and any other

9 methodology that it proposes to use to locate ESI likely to contain responsive and discoverable

10 information. The producing party may provide unique hit counts for each search query.

11   ii.   The requesting party is entitled, within 14 days of the producing

12 party's disclosure, to add no more than 10 search terms or queries to those disclosed by the

13 producing party absent a showing of good cause or written agreement of the parties.

14   iii.   The following provisions apply to search terms / queries of the

15 requesting party. Focused terms and queries should be employed; broad terms or queries, such

16 as product and company names, generally should be avoided. A conjunctive combination of

17 multiple words or phrases (*e.g.*, "computer" and "system") narrows the search and shall count as

18 a single search term. A disjunctive combination of multiple words or phrases (*e.g.*, "computer"

19 or "system") broadens the search, and thus each word or phrase shall count as a separate search

20 term unless they are variants of the same word. The producing party may identify each search

21 term or query returning overbroad results demonstrating the overbroad results and provide a

22 counter proposal correcting the overbroad search or query.

23

24

25 [DRAFT] AGREEMENT REGARDING DISCOVERY OF ELECTRONICALLY
STORED INFORMATION AND [PROPOSED] ORDER   PAGE - 3
26 (Case No. 2:20-CV-203)

1   b.   After production: Within 21 days of the producing party notifying the

2 receiving party that it has substantially completed the production of documents responsive to a

3 request, the responding party may request no more than 10 additional search terms or queries.

4 The immediately preceding section (Section C.2.a.iii) applies.

5   3.   Format.

6   ESI will be produced to the requesting party in accordance with the formatting

7 specifications described in the Civil Division's "Specifications for Production of ESI and

8 Digitized ("Scanned") Images" ("Production Specifications"), attached to this Agreement as

9 Exhibit A subject to the exceptions set forth in subsection C.6 herein.

10   4.   De-duplication. The parties may de-duplicate their ESI production across all

11 custodial and non-custodial data sources after disclosure to the requesting party, and the duplicate

12 custodian information removed during the de-duplication process but tracked in a duplicate/other

13 custodian field in the database load file. Only exact duplicates as tracked by the MD5 hash field

14 will be removed.

15   5.   Email Threading. The parties may use analytics technology to identify email

16 threads and need only produce the unique most inclusive copy and related family members and

17 may exclude lesser inclusive copies. Lesser inclusive versions of an email chain shall be produced

18 when the most inclusive email would not contain an attachment that was attached in a lesser

19 inclusive email. Emails where conversations branch off (such as forwarded emails or when a

20 subset of participants have a concurrent conversation) shall not be considered a lesser inclusive

21 of an email thread. Upon reasonable request, the producing party will produce a less inclusive

22 copy. E-mail threading procedures shall be consistent with section 4 of the Production

23 Specifications.

24

25 [DRAFT] AGREEMENT REGARDING DISCOVERY OF ELECTRONICALLY
STORED INFORMATION AND [PROPOSED] ORDER   PAGE - 4
(Case No. 2:20-CV-203)

26

6.   Metadata fields. The parties agree to produce metadata, to the extent it is reasonably accessible and non-privileged, as described in section 3 of the Production Specifications. Except that the following metadata fields will not be produced: AttachIDs, ConvIndex, Date Time Accd, Date Time Mod, DocDate, Extended properties, Group ID, Message Header, Message Type, and Properties. The parties will, however, produce Beginning and End Attachment IDs.

7.   Hard-Copy Documents. If the parties elect to produce hard-copy documents in an electronic format, they shall follow the procedures for doing so outlined in the Production Specifications.

D.   Preservation of ESI

The parties acknowledge that they have a common law obligation, as expressed in Fed. R. Civ. P. 37(e), to take reasonable and proportional steps to preserve discoverable information in the party's possession, custody, or control. With respect to preservation of ESI, the parties agree as follows:

1.   Absent a showing of good cause by the requesting party, the parties shall not be required to modify the procedures used by them in the ordinary course of business to back-up and archive data; provided, however, that the parties shall preserve all discoverable ESI in their possession, custody, or control.

2.   The parties will supplement their disclosures (including those made pursuant to Section B, *supra*) in accordance with Fed. R. Civ. P. 26(e) with discoverable ESI responsive to a particular discovery request or mandatory disclosure where that data is created after a disclosure or response is made (unless excluded under Sections D.3 or E.1–2).

1   3.   Absent a showing of good cause by the requesting party, the following categories

2   of ESI need not be preserved:

3   a.   Deleted, slack, fragmented, or other data only accessible by forensics.

4   b.   Random access memory (RAM), temporary files, or other ephemeral data that are difficult to preserve without disabling the operating system.

5

        c.   On-line access data such as temporary internet files, history, cache, cookies, and the like.

6

7   d.   Data in metadata fields that are frequently updated automatically, such as last-opened dates (see also Section E.5).

8   e.   Back-up data that are duplicative of data that are more accessible elsewhere.

9

10   f.   Server, system or network logs.

11   g.   Data remaining from systems no longer in use that is unintelligible on the systems in use.

12   h.   Electronic data (*e.g.*, email, calendars, contact data, and notes) sent to or from mobile devices (*e.g.*, iPhone, iPad, Android devices), provided that

13   a copy of all such electronic data is automatically saved in real time elsewhere (such as on a server, laptop, desktop computer, or "cloud"

14   storage).

15   **E.   Privilege**

16   1.   A producing party shall create a privilege log of all documents fully withheld from

17   production on the basis of a privilege or protection, unless otherwise agreed or excepted by this

18   Agreement and Order. Privilege logs shall include a unique identification number for each

19   document and the basis for the claim (attorney-client privileged or work-product protection). For

20   ESI, the privilege log may be generated using available metadata, including author/recipient or

21   to/from/cc/bcc names; the subject matter or title; and date created. Should the available metadata

22   provide insufficient information for the purpose of evaluating the privilege claim asserted, the

23   producing party shall include such additional information as required by the Federal Rules of

24   Civil Procedure. Privilege logs will be produced to all other parties no later than 60 days after

25

26

1   delivering a production unless an earlier deadline is agreed to by the parties.

2   2.   Redactions need not be logged so long as the basis for the redaction is clear on the

3   redacted document.

4   3.   With respect to privileged or work-product information generated after the filing

5   of the complaint, parties are not required to include any such information in privilege logs.

6   4.   Activities undertaken in compliance with the duty to preserve information are

7   protected from disclosure and discovery under Fed. R. Civ. P. 26(b)(3)(A) and (B).

8   5.   Pursuant to Fed. R. Evid. 502(d), the production of any documents in this

9   proceeding shall not, for the purposes of this proceeding or any other federal or state proceeding,

10 constitute a waiver by the producing party of any privilege applicable to those documents,

11 including the attorney-client privilege, attorney work-product protection, or any other privilege

12 or protection recognized by law. Information produced in discovery that is protected as privileged

13 or work product shall be immediately returned to the producing party, and its production shall not

14 constitute a waiver of such protection.

15

16 DATED: _____

17

18 | JOSEPH H. HUNT | HARRIGAN LEYH FARMER & THOMSEN LLP |
    | Assistant Attorney General | |
19 | | By: *s/ Timothy G, Leyh* |
    | BRIAN T. MORAN | By: *s/ Kristin E. Ballinger* |
    | United States Attorney | By: *s/ Shane P. Cramer* |
20 | | Timothy G. Leyh, WSBA #14853 |
    | | Kristin E. Ballinger, WSBA #28253 |
21 | ALEXANDER K. HAAS | Shane P. Cramer, WSBA #35099 |
    | Director | 999 Third Avenue, Suite 4400 |
22 | | Seattle, WA 98104 |
    | JACQUELINE COLEMAN SNEAD | Tel: (206) 623-1700 |
    | Assistant Branch Director | Email: timl@harriganleyh.com |
23 | | Email: kristinb@harriganleyh.com |
    | | Email: shanec@harriganleyh.com |
24

25 [DRAFT] AGREEMENT REGARDING DISCOVERY OF ELECTRONICALLY
    STORED INFORMATION AND [PROPOSED] ORDER   PAGE - 7
26 (Case No. 2:20-CV-203)

1

/s/ Michael J. Gerardi
Michael J. Gerardi

2   (D.C. Bar No. 1017949)
Trial Attorney

3   United States Department of Justice
Civil Division, Federal Programs

4   Branch
1100 L St. NW, Room 12212

5   Washington, D.C. 20005
Tel: (202) 616-0680

6   Fax: (202) 616-8460
E-mail: michael.j.gerardi@usdoj.gov

7
Attorneys for Plaintiff

8

DANIEL T. SATTERBERG
King County Prosecuting Attorney

By: *s/ Timothy P. Barnes*

By: *s/ H. Kevin Wright*
    Timothy P. Barnes, WSBA #29929
    H. Kevin Wright, WSBA #19121
    Senior Deputy Prosecuting Attorneys
    516 Third Avenue, Suite W400
    Seattle, WA 98104
    Tel: (206) 477-1120
    Fax: (206) 296-0191
    Email: timothy.barnes@kingcounty.gov
    Email: kevin.wright@kingcounty.gov

Attorneys for King County and Dow Constantine

9

10

11

12   **CERTIFICATE OF SERVICE**

13   I hereby certify that on June 16, 2020, I served a copy of this [Draft] Agreement

14   Regarding Discovery of Electronically Stored Information and [Proposed] Order on all counsel

15   of record in the case by filing it on the CM / ECF system.

                                          /s/*Michael J. Gerardi*
16   Michael J. Gerardi

17

18

19

20

21

22

23

24

25   [DRAFT] AGREEMENT REGARDING DISCOVERY OF ELECTRONICALLY
      STORED INFORMATION AND [PROPOSED] ORDER   PAGE - 8
26   (Case No. 2:20-CV-203)

1

2

3   **ORDER**

4   Based on the foregoing, IT IS SO ORDERED.

5   DATED: ⸺⸺⸺⸺⸺⸺⸺⸺

6

7   The Honorable <u>Robert J. Bryan</u>
                                             UNITED STATES DISTRICT JUDGE
8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25   [DRAFT] AGREEMENT REGARDING DISCOVERY OF ELECTRONICALLY
     STORED INFORMATION AND [PROPOSED] ORDER   PAGE - 9
26   (Case No. 2:20-CV-203)

EXHIBIT A

September 2017

Specifications for Production of ESI and Digitized ("Scanned") Images ("Production Specifications")

Collection of Electronically Stored Information (ESI)

Careful consideration should be given to the methodology, implementation and documentation of ESI collection to ensure that all responsive data and metadata are preserved in the collection process.

1.   Specification Modifications

Any modifications or deviations from the Production Specifications may be done only with the express permission of the government and these modifications or deviations should be communicated to the government and approved by the government in written form. Any responsive data or documents that exist in locations or native forms not discussed in these Production Specifications remain responsive and, therefore, arrangements should be made with the government to facilitate their production.

2.   Production Format of ESI and Imaged Hard Copy

Responsive ESI and imaged hard copy shall be produced in the format outlined below. All ESI, except as outlined below in sections 5 – 20, shall be rendered to TIFF image format, and accompanied by a Concordance® Image Cross Reference file. All applicable metadata/database (see section 3 below) shall be extracted and provided in Concordance® load file format.

> **a.   Image File Format:** All images, paper documents scanned to images, or rendered ESI, shall be produced as 300 dpi single-page TIFF files, CCITT Group IV (2D Compression). Images should be uniquely and sequentially Bates numbered and unless otherwise specified, Bates numbers should be an endorsement on each image.
>
> > i.   All TIFF file names shall include the unique Bates number burned into the image. (See paragraph 21, below, regarding Bates number instructions.)
> > ii. All TIFF image files shall be stored with the ".tif" extension.
> > iii. Images shall be OCR'd using standard COTS products.
> > > 1.   An exception report shall be provided when limitations of paper digitization software/hardware or attribute conversion do not allow for OCR text conversion of certain images. The report shall include the DOCID or Bates number(s) corresponding to each such image.
> > > > a.   All pages of a document or all pages of a collection of documents that comprise a folder or other logical grouping, including a box, shall be delivered on a single piece of media.
> > > > b.   No image folder shall contain more than 2000 images.
>
> **b.   Concordance® Image Cross Reference file:** Images should be accompanied by a Concordance® Image Cross Reference file that associates each Bates number with its corresponding single-page TIFF image file. The Cross Reference file should also contain the image file path for each Bates numbered page.
>
> > i.   Image Cross Reference Sample Format:
> >
> > ABC00000001,OLS,D:\DatabaseName\Images\001\ ABC00000001.TIF,Y,,,
> > ABC00000002,OLS,D:\DatabaseName\Images\001\ ABC00000002.TIF,,,,
> > ABC00000003,OLS,D:\DatabaseName\Images\001\ ABC00000003.TIF,,,,

1

Specifications for Production of ESI and Digitized ("Scanned") Images
("Production Specifications")

ABC00000004,OLS,D:\DatabaseName\Images\001\ ABC00000004.TIF,Y,,,

c. **Concordance® Load File**: Images should also be accompanied by a "text load file" containing delimited text that will populate fields in a searchable, flat database environment. The file should contain the required fields listed below in section 3.

  i. Text delimited load files are defined using the standard Concordance delimiters. For example:

  Field Separator ¶ *or Code 020*
  Text Qualifier þ *or Code 254*

  ii. This load file should also contain hyperlinks to applicable native files, such as Microsoft Excel or PowerPoint files.
  iii. There should be one line for every record in a collection.
  iv. The load file must contain a header listing the metadata/database fields contained within. For example, if the data file consists of a First Page of a Record (BegDoc#), Last Page of a Record (ending Bates / ENDDOC#), DOCID, DOCDate, File Name, and a Title, then the structure may appear as follows:

  þBEGDOCþ¶þENDDOCþ¶þDOCIDþ¶þDOCDATEþ¶þFILENAMEþ¶þTITLEþ

d. **The extracted/OCR** text should be provided for each document as a separate single text file. The file name should match the BEGDOC# or DOCID for that specific record and be accompanied by the .txt extension.

e. **Directory and folder structure:** The directory structure for productions should be:
  *CaseName***LoadFiles**
  *CaseName***Images** < For supporting images (can include subfolders as needed, should not include more than 2,000 files per folder)
  *CaseName***Natives** <Native Files location (can include subfolders as needed, should not include more than 2,000 files per folder)
  *CaseName***Text** <Extracted Text files location (can include subfolders as needed, should not include more than 2,000 files per folder)

Specifications for Production of ESI and Digitized ("Scanned") Images ("Production Specifications")

3.   Required Metadata/Database Fields

A "✓" denotes that the indicated field should be present in the load file produced. "Other ESI" includes data discussed in sections 5 – 20 below, but does not include email, email repositories (section 11), "stand alone" items (section 12), and imaged hard copy material (section 9). Email, email repositories, and "stand alone" materials (section 12) should comply with "Email" column below. Imaged hard copy materials should comply with the "Hard Copy" column.

Field name	Field Description	Field Type	Field Value	Hard Copy	E-mail	Other ESI
COLLECTION SOURCE	Name of the Company/Organization data was collected from	Text	160	✓	✓	✓
SOURCE ID (BOX #)	Submission/volume/box number	Text	10	✓	✓	✓
CUSTODIAN	Custodian/Source - format: Last, First or ABC Dept.	Text	160	✓	✓	✓
AUTHOR	Creator of the document	Text	500			✓
BEGDOC#	Start Bates (including prefix) - No spaces	Text	60	✓	✓	✓
ENDDOC#	End Bates (including prefix) - No spaces	Text	60	✓	✓	✓
DOCID	Unique document Bates # or populate with the same value as Start Bates (DOCID = BEGDOC#)	Text	60	✓	✓	✓
PGCOUNT	Page Count	Number	10	✓	✓	✓
GROUPID	Contains the Group Identifier for the family, in order to group files with their attachments	Text	60		✓	✓
PARENTID	Contains the Document Identifier of an attachment's parent	Text	60		✓	✓
ATTACHIDS	Child document list; Child DOCID or Child Start Bates	Text – semicolon delimited	Unlimited	✓	✓	✓

3

September 2017

Specifications for Production of ESI and Digitized ("Scanned") Images
("Production Specifications")

Field name	Field Description	Field Type	Field Value	Hard Copy	E-mail	Other ESI
ATTACHLIST	List of Attachment filenames	Text – semicolon delimited	Unlimited		✓	✓
BEGATTACH	Start Bates number of first attachment	Text	60	✓	✓	✓
ENDATTACH	End Bates number of last attachment	Text	60	✓	✓	✓
PROPERTIES	Privilege notations, Redacted, Document Withheld Based On Privilege	Text – semicolon delimited	Unlimited	✓	✓	✓
RECORD TYPE	Use the following choices: Image, Loose E-mail, E-mail, E-Doc, Attachment, Hard Copy or Other. If using Other, please specify what type after Other	Text	60	✓	✓	✓
FROM	Author - format: Last name, First name.	Text	160		✓	✓
TO	Recipient - format: Last name, First name.	Text – semicolon delimited	Unlimited		✓	✓
CC	Carbon Copy Recipients - format: Last name, First name.	Text – semicolon delimited	Unlimited		✓	✓
BCC	Blind Carbon Copy Recipients - format: Last name, First name.	Text – semicolon delimited	Unlimited		✓	✓
SUBJECT	Subject line/Document Title	Text	Unlimited		✓	✓
CONVINDEX	E-mail system ID used to track replies, forwards, etc.	Text	Unlimited		✓	
DOCDATE	Last Modified Date for files and Sent date for e-mail, this field inherits the date for attachments from their parent.	Date	MM/DD/YYYY		✓	✓

4

Specifications for Production of ESI and Digitized ("Scanned") Images ("Production Specifications")

Field name	Field Description	Field Type	Field Value	Hard Copy	E-mail	Other ESI
TEXT FILEPATH	Relative file path of the text file associated with either the extracted text or the OCR	Text	Unlimited	✓	✓	✓
DATE TIME SENT	Date Sent (USE TIME ZONE OF COLLECTION LOCALITY)	Date and Time	MM/DD/YYYY HH:MM:SS		✓	✓
DATE TIME CRTD	Date Created (USE TIME ZONE OF COLLECTION LOCALITY)	Date and Time	MM/DD/YYYY HH:MM:SS		✓	✓
DATE TIME SVD	Date Saved (USE TIME ZONE OF COLLECTION LOCALITY)	Date and Time	MM/DD/YYYY HH:MM:SS		✓	✓
DATE TIME MOD	Date Last Modified (USE TIME ZONE OF COLLECTION LOCALITY)	Date and Time	MM/DD/YYYY HH:MM:SS		✓	✓
DATE TIME RCVD	Date Received (USE TIME ZONE OF COLLECTION LOCALITY)	Date and Time	MM/DD/YYYY HH:MM:SS		✓	
DATE TIME ACCD	Date Accessed (USE TIME ZONE OF COLLECTION LOCALITY)	Date and Time	MM/DD/YYYY HH:MM:SS		✓	✓
FILE SIZE	Native File Size in bytes	Number	10			✓
FILE NAME	File name – name of file as it appeared in its original location	Text	Unlimited			✓
APPLICATION	Application used to create native file (e.g. Excel, Outlook, Word)	Text	160		✓	✓
FILE EXTENSION	Extension for the file (e.g. .doc, .pdf, .wpd)	Text	10		✓	✓
FILEPATH	Data's original source full folder path	Text	Unlimited		✓	✓

Specifications for Production of ESI and Digitized ("Scanned") Images
("Production Specifications")

Field name	Field Description	Field Type	Field Value	Hard Copy	E-mail	Other ESI
NATIVE LINK	Relative file path location to the native file	Text	Unlimited		✓	✓
FOLDER ID	Complete E-mail folder path (e.g. Inbox\Active) or Hard Copy container information (e.g. folder or binder name)	Text	Unlimited	✓	✓	
PARAGRAPH REQUEST NUMBER	Subpoena/request paragraph number to which the document is responsive. Use semicolon to delimit multiple entries.	Text – semicolon delimited	Unlimited	✓	✓	✓
MD5 HASH	MD5 Hash value (used for deduplication or other processing) (e-mail hash values must be run with the e-mail and all of its attachments)	Text	Unlimited		✓	✓
MESSAGEHEADER	E-mail header. Can contain IP address	Text	Unlimited		✓	
ATTACHMCOUNT	Number of attachments (any level child document) associated with a ParentID	Text	10		✓	
FILE TYPE	Identifies the application that created the file Type of file, not to be confused with file extension	Text	160		✓	✓
COMMENTS	Identifies whether the document has comments associated with it	Text	10		✓	✓
MESSAGE TYPE	Exchange Message class or equivalent	Text	60		✓	
EXTENDED PROPERTIES	For PDFs Only	Text	600		✓	✓

6

Specifications for Production of ESI and Digitized ("Scanned") Images ("Production Specifications")

4.   Search, De-Duplication, Near-Duplicate Identification, E-mail Conversation Threading and Other Culling Procedures

De-duplication of exact copies <u>within</u> a custodian's data may be done at the family level, but all file paths and custodians must be provided for each duplicate document in an exception report in .csv format. The recipient shall not use any other procedure to cull, filter, group, separate or de-duplicate, or near-deduplicate, etc. (i.e., reduce the volume of) responsive material before discussing with and obtaining the written approval of the government. All objective coding (e.g., near duplicate ID or e-mail thread ID) shall be discussed and produced to the government as additional metadata fields. The recipient will not employ analytic software or technology to search identify, or review potentially responsive material, including but not limited to technology assisted review (TAR) or predictive coding, without first discussing with the government.

5.   Hidden Text

All hidden text (e.g. track changes, hidden columns, mark-ups, notes) shall be expanded and rendered in the image file. For files that cannot be expanded the native files shall be produced with the image file.

6.   Embedded Files

All non-graphic embedded objects (Word documents, Excel spreadsheets, .wav files, etc.) that are found within a file shall be extracted and produced. For purposes of production, the embedded files shall be treated as attachments to the original file, with the parent/child relationship preserved.

7.   Image-Only Files

All image-only files (non-searchable .pdfs, multi-page TIFFs, Snipping Tool and other screenshots, etc., as well as all other images that contain text) shall be produced with OCR text and metadata/database fields identified in section 3 for "Other ESI."

8.   Encrypted Files

Any data (whether individual files or digital containers) that is protected by a password, encryption key, digital rights management, or other encryption scheme, shall be decrypted prior to processing for production.

> **a.**   The unencrypted text shall be extracted and provided per section 2.c. The unencrypted files shall be used to render images and provided per sections 2.a and 2.b. The unencrypted native file shall be produced pursuant to sections 10-20.
> **b.**   If such protected data is encountered but unable to be processed, each file or container shall be reported as an exception in the accompanying Exception Report (pursuant to section 26) and shall include all available metadata associated with the data, including custodian information.

9.   Production of Imaged Hard Copy Records

All imaged hard copy material shall reflect accurate document unitization including all attachments and container information (to be reflected in the PARENTID, ATTACHID, BEGATTACH, ENDATTACH and FOLDERID).

**Specifications for Production of ESI and Digitized ("Scanned") Images
("Production Specifications")**

     a.   Unitization in this context refers to identifying and marking the boundaries of documents within the collection, where a document is defined as the smallest physical fastened unit within a bundle. (e.g., staples, paperclips, rubber bands, folders, or tabs in a binder).

     b.   The first document in the collection represents the parent document and all other documents will represent the children.

     c.   All imaged hard copy documents shall be produced as 300 dpi single-page TIFF files, CCITT Group IV (2D Compression). All documents shall be produced in black and white TIFF format unless the image requires color. An image requires color when color in the document adds emphasis to information in the document or is itself information that would not be readily apparent on the face of a black and white image. Images identified as requiring color shall be produced as color 300 dpi single-page JPEG files.

     d.   All objective coding (e.g., document date or document author) should be discussed and produced to the government as additional metadata/database fields.

10.   **Production of Spreadsheets and Presentation Files**

All spreadsheet and presentation files (e.g. Excel, PowerPoint) shall be produced in the unprocessed "as kept in the ordinary course of business" state (i.e., in native format), with an associated placeholder image and endorsed with a unique Bates number. *See* section 20 below. The file produced should maintain the integrity of all source, custodian, application, embedded and related file system metadata. No alteration shall be made to file names or extensions for responsive native electronic files.

11.   **Production of E-mail Repositories**

E-mail repositories, also known as e-mail databases (e.g., Outlook PST, Lotus NSF, etc.), can contain a variety of items, including: messages, calendars, contacts, tasks, etc. For purposes of production, responsive items shall include the "E-mail" metadata/database fields outlined in section 3, including but not limited to all parent items (mail, calendar, contacts, tasks, notes, etc.) and child files (attachments of files to e-mail or other items) with the parent/child relationship preserved. Our preferred format for e-mail productions is PST. E-mail should NOT be provided in X400 or X500 format. E-mail databases from systems other than Microsoft Exchange shall be produced after consultation with and written consent of the government about the format for the production of such databases.

12.   **Production of Items Originally Generated in E-mail Repositories but Found and Collected Outside of E-mail Repositories, i.e., "Stand-alone" Items**

Any parent e-mail or other parent items (e.g., calendar, contacts, tasks, notes, etc.) found and collected outside of e-mail repositories (e.g., items having extensions .msg, .htm, .mht, etc.), shall be produced with the "Loose E-mail" metadata fields outlined in section 3, including but not limited to any attachments, maintaining the family (parent/child) relationship.

13.   **Production of Instant Messenger (IM), Voicemail Data, Audio Data, Video Data, etc.**

The responding party shall identify, collect, and produce any and all data which is responsive to the requests which may be stored in audio or video recordings, cell phone/PDA/Blackberry/smart phone data, tablet data, voicemail messaging data, instant messaging, text messaging, conference call data, video/audio conferencing (e.g., GoTo Meeting, WebEx), and related/similar technologies. However, such data, logs, metadata or other files related thereto, as well as other

8

**Specifications for Production of ESI and Digitized ("Scanned") Images
("Production Specifications")**

less common but similar data types, shall be produced after consultation with and written consent of the government about the format for the production of such data.

14.   Production of Social Media

Prior to any production of responsive data from social media (e.g., Twitter, Facebook, Google+, LinkedIn, etc.) the producing party shall first discuss with the government the potential export formats before collecting the information, to ensure it is collected and produced in a way that preserves the original metadata, has a clear chain of custody, and provides as much information as possible regarding the source and history of each individual communication.

Social media platforms offer different functions, forms of content, and capability for downloading accounts. Because of these differences, prior to collection of social media data, the producing party must discuss with the government the available export and production methods and formats that the producing party is considering. Unless the government agrees to an alternative in writing, regardless of the social media platform, productions of social media content must meet the following general requirements: (1) separate (2) searchable (3) static images of (4) each responsive posting on the social media platform, (5) all related content (e.g., comments, likes, share or re-transmittal information, images, videos, linked documents and content), and (6) associated metadata (e.g., user name(s), date, and time of all posts, comments, likes, share or re-transmittals).

These general requirements are in addition to any more specific requirements in a particular request (e.g., geolocation data), and the producing party must ask the government about any perceived conflict between these requirements and another source of specifications or requirements. If available from the social media platform or through social media data processing software, files that facilitate interactive review of the data (i.e., html files) as well as load files in .csv format must be produced with the associated content.

15.   Production of Structured Data

Prior to any production of responsive data from a structured database (e.g., Oracle, SAP, SQL, MySQL, QuickBooks, etc.), the producing party shall first identify the database type and version number, provide the database dictionary and any user manuals, or any other documentation describing the structure and/or content of the database and a list of all reports that can be generated from the database. The list of reports shall be provided in native Excel (.xls or .xlsx) format.

16.   Production of Structured Data from Proprietary Applications

Prior to any production of structured data from proprietary applications (e.g., proprietary timekeeping, accounting, sales rep call notes, CRMs, SharePoint etc.) the producing party shall first provide the database dictionary and a list of all reports that can be generated from the structured database. The list of reports shall be produced in native Excel (.xls or .xlsx) format.

17.   Production of Photographs with Native File or Digitized ESI

Photographs shall be produced as single-page JPEG files with a resolution equivalent to the original image as they were captured/created. All JPEG files shall have extracted metadata/database fields provided in a Concordance® load file format as outlined in section 3 for "Other ESI."

September 2017

Specifications for Production of ESI and Digitized ("Scanned") Images
("Production Specifications")

18.   Production of Images from which Text Cannot be OCR Converted

An exception report shall be provided when limitations of paper digitization software/hardware or attribute conversion do not allow for OCR text conversion of certain images. The report shall include the DOCID or Bates number(s) corresponding to each such image.

19.   Production of ESI from Non-PC or Non-Windows-based Systems

If responsive ESI is in non-PC or non-Windows-based Systems (e.g., Apple, IBM mainframes, and UNIX machines, Android device, etc.), the ESI shall be produced after discussion with and written consent of the government about the format for the production of such data.

20.   Production of Native Files (When Applicable Pursuant to These Specifications)

Production of native files, as called for in these specifications, shall have extracted metadata/database fields provided in a Concordance® load file format as defined in the field specifications for "Other ESI" as outlined in section 3.

ESI shall be produced in a manner which is functionally usable by the government. The following are examples:

> **a.**   AutoCAD data, e.g., DWG and DXF files, shall be processed/converted and produced as single-page JPG image files and accompanied by a Concordance® Image formatted load file as described above. The native files shall be placed in a separate folder on the production media and linked by a hyperlink within the text load file.
> **b.**   GIS data shall be produced in its native format and be accompanied by a viewer such that the mapping or other data can be reviewed in a manner that does not detract from its ability to be reasonably understood.
> **c.**   Audio and video recordings shall be produced in native format and be accompanied by a viewer if such recordings do not play in a generic application (e.g., Windows Media Player).

21.   Bates Number Convention

All images should be assigned Bates numbers before production to the government. Each Bates number shall be a standard length, include leading zeros in the number, and be unique for each produced page. The numbers should be endorsed on the actual images at a location that does not obliterate, conceal, or interfere with any information from the source document. Native files should be assigned a single Bates number for the entire file which will represent the native document in the load/dat file. The load/dat file will include a reference to the native file path and utilize the NATIVELINK metadata field). The Bates number shall not exceed 30 characters in length and shall include leading zeros in the numeric portion. The Bates number shall be a unique number given to each page (when assigned to an image) or to each document (when assigned to a native file). If the parties agree to a rolling production, the numbering convention shall remain consistent throughout the entire production. There shall be no spaces between the prefix and numeric value. If suffixes are required, please use "dot notation." Below is a sample of dot notation:

> PREFIX0000001   PREFIX0000003
> PREFIX0000001.001   PREFIX0000003.001
> PREFIX0000001.002   PREFIX0000003.002

10

September 2017

Specifications for Production of ESI and Digitized ("Scanned") Images
("Production Specifications")

---

22.   Media Formats for Storage and Delivery of Production Data

Electronic documents and data shall be delivered on any of the following media:

   a.   CD-ROMs and/or DVD-R (+/-) formatted to ISO/IEC 13346 and Universal Disk
   Format 1.02 specifications; Blu-ray.
   b.   External hard drives (USB 3.0 or higher, Firewire or eSATA, formatted to NTFS
   format specifications) or flash drives.
   c.   Storage media used to deliver ESI shall be appropriate to the size of the data in
   the production.
   d.   Media should be labeled with the case name, production date, Bates range, and
   producing party.

23.   Virus Protection and Security for Delivery of Production Data

Production data shall be free of computer viruses. Any files found to include a virus shall be
quarantined by the producing party and noted in a log to be provided to the government.
Password protected or encrypted files or media shall be provided with corresponding passwords
and specific decryption instructions. No encryption software shall be used without the written
consent of the government.

24.   Compliance and Adherence to Generally Accepted Technical Standards

Production shall be in conformance with standards and practices established by the National
Institute of Standards and Technology ("NIST" at www.nist.gov), U.S. National Archives &
Records Administration ("NARA" at www.archives.gov), American Records Management
Association ("ARMA International" at www.arma.org), American National Standards Institute
("ANSI" at www.ansi.org), International Organization for Standardization ("ISO" at
www.iso.org), and/or other U.S. Government or professional organizations.

25.   Read Me Text File

All deliverables shall include a "read me" text file at the root directory containing: total number
of records, total number of images/pages or files, mapping of fields to plainly identify field
names, types, lengths, and formats. The file shall also indicate the field name to which images
will be linked for viewing, date and time format, and confirmation that the number of files in load
files matches the number of files produced.

26.   Exception Report

An exception report, in .csv format, shall be included, documenting any production anomalies
utilizing the electronic Bates number (DOCID or control numbering) assigned during the
collection, processing, and production phases.

27.   Transmittal Letter to Accompany Deliverables

All deliverables should be accompanied by a transmittal letter including the production date, case
name and number, producing party name, and Bates number range produced.

-XXX-

11